Fengzhu Sun

List of Publications by Year
in descending order

Source: https://exaly.com/author-pdf/6487826/publications.pdf

Version: 2024-02-01

		36303	38395
185	10,709	51	95
papers	citations	h-index	g-index

202	202	202	12512
all docs	docs citations	times ranked	citing authors

#	Article	IF	Citations
1	Correlation detection strategies in microbial data sets vary widely in sensitivity and precision. ISME Journal, 2016, 10, 1669-1681.	9.8	593
2	Marine bacterial, archaeal and protistan association networks reveal ecological linkages. ISME Journal, 2011, 5, 1414-1425.	9.8	560
3	VirFinder: a novel k-mer based tool for identifying viral sequences from assembled metagenomic data. Microbiome, 2017, 5, 69.	11.1	433
4	Inferring Domain-Domain Interactions From Protein-Protein Interactions. Genome Research, 2002, 12, 1540-1548.	5.5	376
5	A dynamic programming algorithm for haplotype block partitioning. Proceedings of the National Academy of Sciences of the United States of America, 2002, 99, 7335-7339.	7.1	312
6	Identifying viruses from metagenomic data using deep learning. Quantitative Biology, 2020, 8, 64-77.	0.5	302
7	Local similarity analysis reveals unique associations among marine bacterioplankton species and environmental factors. Bioinformatics, 2006, 22, 2532-2538.	4.1	292
8	Prediction of Protein Function Using Proteinâ€“Protein Interaction Data. Journal of Computational Biology, 2003, 10, 947-960.	1.6	274
9	Alignment-free d_2^* oligonucleotide frequency dissimilarity measure improves prediction of hosts from metagenomically-derived viral sequences. Nucleic Acids Research, 2017, 45, 39-53.	14.5	245
10	Haplotype Block Structure and Its Applications to Association Studies: Power and Study Designs. American Journal of Human Genetics, 2002, 71, 1386-1394.	6.2	243
11	Extended local similarity analysis (eLSA) of microbial community and other time series data with replicates. BMC Systems Biology, 2011, 5, S15.	3.0	223
12	Taq DNA polymerase slippage mutation rates measured by PCR and quasi-likelihood analysis: (CA/GT)n and (A/T)n microsatellites. Nucleic Acids Research, 2003, 31, 974-980.	14.5	221
13	A critical assessment of Mus musculus gene function prediction using integrated genomic evidence. Genome Biology, 2008, 9, S2.	9.6	214
14	CancerLocator: non-invasive cancer diagnosis and tissue-of-origin prediction using methylation profiles of cell-free DNA. Genome Biology, 2017, 18, 53.	8.8	204
15	The Relationship Between Microsatellite Slippage Mutation Rate and the Number of Repeat Units. Molecular Biology and Evolution, 2003, 20, 2123-2131.	8.9	195
16	Alignment-Free Sequence Comparison (I): Statistics and Power. Journal of Computational Biology, 2009, 16, 1615-1634.	1.6	187
17	Transmission/Disequilibrium Tests Using Multiple Tightly Linked Markers. American Journal of Human Genetics, 2000, 67, 936-946.	6.2	159
18	Mapping gene ontology to proteins based on protein-protein interaction data. Bioinformatics, 2004, 20, 895-902.	4.1	149

#	Article	IF	Citations
19	Benchmarking of alignment-free sequence comparison methods. Genome Biology, 2019, 20, 144.	8.8	147
20	Haplotype Block Partitioning and Tag SNP Selection Using Genotype Data and Their Applications to Association Studies. Genome Research, 2004, 14, 908-916.	5.5	143
21	Efficient statistical significance approximation for local similarity analysis of high-throughput time series data. Bioinformatics, 2013, 29, 230-237.	4.1	137
22	GOLabeler: improving sequence-based large-scale protein function prediction by learning to rank. Bioinformatics, 2018, 34, 2465-2473.	4.1	136
23	An Integrated Probabilistic Model for Functional Prediction of Proteins. Journal of Computational Biology, 2004, 11, 463-475.	1.6	134
24	Critical Assessment of Metagenome Interpretation: the second round of challenges. Nature Methods, 2022, 19, 429-440.	19.0	133
25	COCACOLA: binning metagenomic contigs using sequence COmposition, read CoverAge, CO-alignment and paired-end read LinkAge. Bioinformatics, 2017, 33, 791-798.	4.1	130
26	Case-Only Design to Measure Gene-Gene Interaction. Epidemiology, 1999, 10, 167-170.	2.7	127
27	Further understanding human disease genes by comparing with housekeeping genes and other genes. BMC Genomics, 2006, 7, 31.	2.8	126
28	Variance adjusted weighted UniFrac: a powerful beta diversity measure for comparing communities based on phylogeny. BMC Bioinformatics, 2011, 12, 118.	2.6	126
29	New developments of alignment-free sequence comparison: measures, statistics and next-generation sequencing. Briefings in Bioinformatics, 2014, 15, 343-353.	6.5	126
30	Association of relative brain age with tobacco smoking, alcohol consumption, and genetic variants. Scientific Reports, 2020, 10, 10.	3.3	121
31	HapBlock: haplotype block partitioning and tag SNP selection software using a set of dynamic programming algorithms. Bioinformatics, 2005, 21, 131-134.	4.1	109
32	CGI: a new approach for prioritizing genes by combining gene expression and proteinâ€“protein interaction data. Bioinformatics, 2007, 23, 215-221.	4.1	106
33	DiseaseConnect: a comprehensive web server for mechanism-based diseaseâ€“disease connections. Nucleic Acids Research, 2014, 42, W137-W146.	14.5	106
34	Alignment-Free Sequence Comparison (II): Theoretical Power of Comparison Statistics. Journal of Computational Biology, 2010, 17, 1467-1490.	1.6	105
35	Accurate Genome Relative Abundance Estimation Based on Shotgun Metagenomic Reads. PLoS ONE, 2011, 6, e27992.	2.5	105
36	Cross-depth analysis of marine bacterial networks suggests downward propagation of temporal changes. ISME Journal, 2015, 9, 2573-2586.	9.8	105

#	Article	IF	Citations
37	Transmission Disequilibrium Test (TDT) when Only One Parent Is Available The 1-TDT. American Journal of Epidemiology, 1999, 150, 97-104.	3.4	104
38	Spatiotemporal profile of postsynaptic interactomes integrates components of complex brain disorders. Nature Neuroscience, 2017, 20, 1150-1161.	14.8	104
39	An integrated approach to the prediction of domain-domain interactions. BMC Bioinformatics, 2006, 7, 269.	2.6	102
40	HAPLORE: a program for haplotype reconstruction in general pedigrees without recombination. Bioinformatics, 2005, 21, 90-103.	4.1	96
41	Diffusion Kernel-Based Logistic Regression Models for Protein Function Prediction. OMICS A Journal of Integrative Biology, 2006, 10, 40-55.	2.0	90
42	An integrative approach for causal gene identification and gene regulatory pathway inference. Bioinformatics, 2006, 22, e489-e496.	4.1	87
43	NetGO: improving large-scale protein function prediction with massive network information. Nucleic Acids Research, 2019, 47, W379-W387.	14.5	86
44	Alignment-Free Sequence Analysis and Applications. Annual Review of Biomedical Data Science, 2018, 1, 93-114.	6.5	78
45	Alignment-Free Sequence Comparison Based on Next-Generation Sequencing Reads. Journal of Computational Biology, 2013, 20, 64-79.	1.6	76
46	Comparison of metagenomic samples using sequence signatures. BMC Genomics, 2012, 13, 730.	2.8	74
47	Gut microbial and metabolomic profiles after fecal microbiota transplantation in pediatric ulcerative colitis patients. FEMS Microbiology Ecology, 2018, 94, .	2.7	73
48	Inferring activity changes of transcription factors by binding association with sorted expression profiles. BMC Bioinformatics, 2007, 8, 452.	2.6	72
49	A network-based integrated framework for predicting virusâ€“prokaryote interactions. NAR Genomics and Bioinformatics, 2020, 2, lqaa044.	3.2	69
50	The Polymerase Chain Reaction and Branching Processes. Journal of Computational Biology, 1995, 2, 63-86.	1.6	68
51	CAFE: aCcelerated Alignment-FrEe sequence analysis. Nucleic Acids Research, 2017, 45, W554-W559.	14.5	59
52	Ecdysone Receptor Acts in fruitless- Expressing Neurons to Mediate Drosophila Courtship Behaviors. Current Biology, 2009, 19, 1447-1452.	3.9	57
53	Mitochondrial genome mutations in hypertensive individuals*1. American Journal of Hypertension, 2004, 17, 629-635.	2.0	56
54	SolidBin: improving metagenome binning with semi-supervised normalized cut. Bioinformatics, 2019, 35, 4229-4238.	4.1	52

#	Article	IF	Citations
55	Somatic, germline and sex hierarchy regulated gene expression during Drosophila metamorphosis. BMC Genomics, 2009, 10, 80.	2.8	49
56	Alignment-free protein interaction network comparison. Bioinformatics, 2014, 30, i430-i437.	4.1	48
57	Classifying Alzheimer's disease with brain imaging and genetic data using a neural network framework. Neurobiology of Aging, 2018, 68, 151-158.	3.1	48
58	Haplotype Block Partition with Limited Resources and Applications to Human Chromosome 21 Haplotype Data. American Journal of Human Genetics, 2003, 73, 63-73.	6.2	47
59	Maternal influence on blood pressure suggests involvement of mitochondrial DNA in the pathogenesis of hypertension: the Framingham Heart Study. Journal of Hypertension, 2007, 25, 2067-2073.	0.5	47
60	Network motif identification in stochastic networks. Proceedings of the National Academy of Sciences of the United States of America, 2006, 103, 9404-9409.	7.1	45
61	New powerful statistics for alignment-free sequence comparison under a pattern transfer model. Journal of Theoretical Biology, 2011, 284, 106-116.	1.7	43
62	Sequence-Based Prioritization of Nonsynonymous Single-Nucleotide Polymorphisms for the Study of Disease Mutations. American Journal of Human Genetics, 2007, 81, 346-360.	6.2	42
63	Parity is associated with cognitive function and brain age in both females and males. Scientific Reports, 2020, 10, 6100.	3.3	41
64	Defining haplotype blocks and tag single-nucleotide polymorphisms in the human genome. Human Molecular Genetics, 2004, 13, 335-342.	2.9	39
65	Assessment of the reliability of protein-protein interactions and protein function prediction. Pacific Symposium on Biocomputing Pacific Symposium on Biocomputing, 2003, , 140-51.	0.7	39
66	Association of Genetic Variation in the Mitochondrial Genome With Blood Pressure and Metabolic Traits. Hypertension, 2012, 60, 949-956.	2.7	38
67	Prediction of virus-host infectious association by supervised learning methods. BMC Bioinformatics, 2017, 18, 60.	2.6	38
68	A network-based integrative approach to prioritize reliable hits from multiple genome-wide RNAi screens in Drosophila. BMC Genomics, 2009, 10, 220.	2.8	37
69	A Novel Class of Tests for the Detection of Mitochondrial DNAâ€“Mutation Involvement in Diseases. American Journal of Human Genetics, 2003, 72, 1515-1526.	6.2	36
70	Integrating multiple protein-protein interaction networks to prioritize disease genes: a Bayesian regression approach. BMC Bioinformatics, 2011, 12, S11.	2.6	36
71	A unified approach for allele frequency estimation, SNP detection and association studies based on pooled sequencing data using EM algorithms. BMC Genomics, 2013, 14, S1.	2.8	32
72	Prediction of protein function using protein-protein interaction data. , 0, , .		31

#	Article	IF	Citations
73	A global overview of genetically interpretable multimorbidities among common diseases in the UK Biobank. Genome Medicine, 2021, 13, 110.	8.2	31
74	Whole genome amplification of single cells: mathematical analysis of PEP and tagged PCR. Nucleic Acids Research, 1995, 23, 3034-3040.	14.5	29
75	A Mathematical Analysis ofin VitroMolecular Selection â€“ Amplification. Journal of Molecular Biology, 1996, 258, 650-660.	4.2	29
76	Estimation of the Mutation Rate During Error-prone Polymerase Chain Reaction. Journal of Computational Biology, 2000, 7, 143-158.	1.6	29
77	Alignment-free Transcriptomic and Metatranscriptomic Comparison Using Sequencing Signatures with Variable Length Markov Chains. Scientific Reports, 2016, 6, 37243.	3.3	29
78	MicroPro: using metagenomic unmapped reads to provide insights into human microbiota and disease associations. Genome Biology, 2019, 20, 154.	8.8	29
79	Genomic mapping by end-characterized random clones: a mathematical analysis. Genomics, 1995, 26, 84-100.	2.9	25
80	Searching for interpretable rules for disease mutations: a simulated annealing bump hunting strategy. BMC Bioinformatics, 2006, 7, 417.	2.6	25
81	Modeling RNA degradation for RNA-Seq with applications. Biostatistics, 2012, 13, 734-747.	1.5	25
82	Inference of Markovian properties of molecular sequences from NGS data and applications to comparative genomics. Bioinformatics, 2016, 32, 993-1000.	4.1	25
83	Comparative studies of alignment, alignment-free and SVM based approaches for predicting the hosts of viruses based on viral sequences. Scientific Reports, 2018, 8, 10032.	3.3	25
84	Modeling DNA Shuffling. Journal of Computational Biology, 1999, 6, 77-90.	1.6	23
85	Test of Association for Quantitative Traits in General Pedigrees: The Quantitative Pedigree Disequilibrium Test. Genetic Epidemiology, 2001, 21, S370-5.	1.3	23
86	Extreme Value Distribution Based Gene Selection Criteria for Discriminant Microarray Data Analysis Using Logistic Regression. Journal of Computational Biology, 2004, 11, 215-226.	1.6	23
87	A model-based approach to selection of tag SNPs. BMC Bioinformatics, 2006, 7, 303.	2.6	23
88	Comparison of Metatranscriptomic Samples Based on k-Tuple Frequencies. PLoS ONE, 2014, 9, e84348.	2.5	22
89	Afann: bias adjustment for alignment-free sequence comparison based on sequencing data using neural network regression. Genome Biology, 2019, 20, 266.	8.8	22
90	16S rRNA and metagenomic shotgun sequencing data revealed consistent patterns of gut microbiome signature in pediatric ulcerative colitis. Scientific Reports, 2022, 12, 6421.	3.3	22

#	Article	IF	Citations
91	DomainRBF: a Bayesian regression approach to the prioritization of candidate domains for complex diseases. BMC Systems Biology, 2011, 5, 55.	3.0	21
92	Integrative approaches for predicting protein function and prioritizing genes for complex phenotypes using protein interaction networks. Briefings in Bioinformatics, 2014, 15, 685-698.	6.5	21
93	HiCBin: binning metagenomic contigs and recovering metagenome-assembled genomes using Hi-C contact maps. Genome Biology, 2022, 23, 63.	8.8	20
94	A dynamic programming algorithm for binning microbial community profiles. Bioinformatics, 2006, 22, 1508-1514.	4.1	18
95	Testing gene set enrichment for subset of genes: Sub-GSE. BMC Bioinformatics, 2008, 9, 362.	2.6	18
96	Improving contig binning of metagenomic data using $$ {d}_2^S $$ oligonucleotide frequency dissimilarity. BMC Bioinformatics, 2017, 18, 425.	2.6	18
97	Assessment of metagenomic assemblers based on hybrid reads of real and simulated metagenomic sequences. Briefings in Bioinformatics, 2020, 21, 777-790.	6.5	18
98	CEDER: Accurate Detection of Differentially Expressed Genes by Combining Significance of Exons Using RNA-Seq. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2012, 9, 1281-1292.	3.0	17
99	Identifying Group-Specific Sequences for Microbial Communities Using Long k-mer Sequence Signatures. Frontiers in Microbiology, 2018, 9, 872.	3.5	17
100	Reads Binning Improves Alignment-Free Metagenome Comparison. Frontiers in Genetics, 2019, 10, 1156.	2.3	17
101	On the use of population-based registries in the clinical validation of genetic tests for disease susceptibility. Genetics in Medicine, 2000, 2, 186-192.	2.4	16
102	The Mutation Process of Microsatellites During the Polymerase Chain Reaction. Journal of Computational Biology, 2003, 10, 143-155.	1.6	16
103	Inference of domain-disease associations from domain-protein, protein-disease and disease-disease relationships. BMC Systems Biology, 2016, 10, 4.	3.0	16
104	Improving brain age estimates with deep learning leads to identification of novel genetic factors associated with brain aging. Neurobiology of Aging, 2021, 105, 199-204.	3.1	16
105	Metagenomic Analyses of Multiple Gut Datasets Revealed the Association of Phage Signatures in Colorectal Cancer. Frontiers in Cellular and Infection Microbiology, 0, 12, .	3.9	16
106	The effects of protein interactions, gene essentiality and regulatory regions on expression variation. BMC Systems Biology, 2008, 2, 54.	3.0	15
107	Compound Poisson Approximation of the Number of Occurrences of a Position Frequency Matrix (PFM) on Both Strands. Journal of Computational Biology, 2008, 15, 547-564.	1.6	14
108	Next-generation sequencing revealed divergence in deletions of the preS region in the HBV genome between different HBV-related liver diseases. Journal of General Virology, 2017, 98, 2748-2758.	2.9	14

#	Article	IF	Citations
109	Increasing prediction performance of colorectal cancer disease status using random forests classification based on metagenomic shotgun sequencing data. Synthetic and Systems Biotechnology, 2022, 7, 574-585.	3.7	14
110	Detecting susceptibility genes in case-control studies using set association. BMC Genetics, 2003, 4, S9.	2.7	13
111	The transmission disequilibrium test and imprinting effects test based on caseâ€parent pairs. Genetic Epidemiology, 2007, 31, 273-287.	1.3	13
112	Statistical significance approximation in local trend analysis of high-throughput time-series data using the theory of Markov chains. BMC Bioinformatics, 2015, 16, 301.	2.6	13
113	Multilevel regularized regression for simultaneous taxa selection and network construction with metagenomic count data. Bioinformatics, 2015, 31, 1067-1074.	4.1	13
114	The Power of Detecting Enriched Patterns: An HMM Approach. Journal of Computational Biology, 2010, 17, 581-592.	1.6	12
115	Detection of Parent-of-Origin Effects for Quantitative Traits in Complete and Incomplete Nuclear Families With Multiple Children. American Journal of Epidemiology, 2011, 174, 226-233.	3.4	12
116	Whole genome association studies for genes affecting alcohol dependence. Genetic Epidemiology, 1999, 17, S337-42.	1.3	11
117	The Power of Transmission Disequilibrium Tests for Quantitative Traits. Genetic Epidemiology, 2001, 21, S632-7.	1.3	11
118	Multiple alignment-free sequence comparison. Bioinformatics, 2013, 29, 2690-2698.	4.1	11
119	Sparse generalized linear model with L 0 approximation for feature selection and prediction with big omics data. BioData Mining, 2017, 10, 39.	4.0	11
120	Prediction of protein function using protein-protein interaction data. Proceedings, 2002, 1, 197-206.	0.1	11
121	Sample sizes for the transmission disequilibrium tests: tdt, s-tdt and 1-tdt. Communications in Statistics - Theory and Methods, 2000, 29, 1129-1142.	1.0	10
122	Assessing the power of tag SNPs in the mapping of quantitative trait loci (QTL) with extremal and random samples. , 2005, 6, 51.		10
123	Background Adjusted Alignment-Free Dissimilarity Measures Improve the Detection of Horizontal Gene Transfer. Frontiers in Microbiology, 2018, 9, 711.	3.5	10
124	MARD: a new method to detect differential gene expression in treatment-control time courses. Bioinformatics, 2006, 22, 2650-2657.	4.1	9
125	Normalizing Metagenomic Hi-C Data and Detecting Spurious Contacts Using Zero-Inflated Negative Binomial Regression. Journal of Computational Biology, 2022, 29, 106-120.	1.6	9
126	Microsatellite mutations during the polymerase chain reaction: mean field approximations and their applications. Journal of Theoretical Biology, 2003, 224, 127-137.	1.7	8

#	Article	IF	Citations
127	Sequence Alignment as Hypothesis Testing. Journal of Computational Biology, 2011, 18, 677-691.	1.6	8
128	Optimal choice of word length when comparing two Markov sequences using a Ï‡ 2-statistic. BMC Genomics, 2017, 18, 732.	2.8	8
129	DeepLINK: Deep learning inference using knockoffs with applications to genomics. Proceedings of the National Academy of Sciences of the United States of America, 2021, 118, .	7.1	8
130	Conservation and implications of eukaryote transcriptional regulatory regions across multiple species. BMC Genomics, 2008, 9, 623.	2.8	7
131	Deep sequencing of HBV pre-S region reveals high heterogeneity of HBV genotypes and associations of word pattern frequencies with HCC. PLoS Genetics, 2018, 14, e1007206.	3.5	7
132	KmerGO: A Tool to Identify Group-Specific Sequences With k-mers. Frontiers in Microbiology, 2020, 11, 2067.	3.5	7
133	AC-PCoA: Adjustment for confounding factors using principal coordinate analysis. PLoS Computational Biology, 2022, 18, e1010184.	3.2	7
134	Normal and Compound Poisson Approximations for Pattern Occurrences in NGS Reads. Journal of Computational Biology, 2012, 19, 839-854.	1.6	6
135	Alignment-free genome comparison enables accurate geographic sourcing of white oak DNA. BMC Genomics, 2018, 19, 896.	2.8	6
136	Statistical significance approximation for local similarity analysis of dependent time series data. BMC Bioinformatics, 2019, 20, 53.	2.6	6
137	Rapid diagnosis and comprehensive bacteria profiling of sepsis based on cell-free DNA. Journal of Translational Medicine, 2020, 18, 5.	4.4	6
138	ARG-SHINE: improve antibiotic resistance class prediction by integrating sequence homology, functional information and deep convolutional neural network. NAR Genomics and Bioinformatics, 2021, 3, lqab066.	3.2	6
139	Decreased secondary faecal bile acids in children with ulcerative colitis and Clostridioides difficile infection. Alimentary Pharmacology and Therapeutics, 2021, 54, 792-804.	3.7	6
140	Finding Genetic Overlaps Among Diseases Based on Ranked Gene Lists. Journal of Computational Biology, 2015, 22, 111-123.	1.6	5
141	Testing for contributions of mitochondrial DNA mutations to complex diseases. , 1998, 15, 451-469.		4
142	Sampling distribution for microsatellites amplified by PCR: mean field approximation and its applications to genotyping. Journal of Theoretical Biology, 2004, 228, 185-194.	1.7	4
143	Prioritizing functional modules mediating genetic perturbations and their phenotypic effects: a global strategy. Genome Biology, 2008, 9, R174.	9.6	4
144	A new statistic for efficient detection of repetitive sequences. Bioinformatics, 2019, 35, 4596-4606.	4.1	4

#	Article	IF	Citations
145	Bayesian Models and Gibbs Sampling Strategies for Local Graph Alignment and Motif Identification in Stochastic Biological Networks. Communications in Information and Systems, 2009, 9, 347-370.	0.5	4
146	The Machine-Learning-Mediated Interface of Microbiome and Genetic Risk Stratification in Neuroblastoma Reveals Molecular Pathways Related to Patient Survival. Cancers, 2022, 14, 2874.	3.7	4
147	Whole Genome Amplification and Branching Processes. Advances in Applied Probability, 1997, 29, 629-668.	0.7	3
148	The method of sib-pair linkage analysis in context of case-control design. Genetic Epidemiology, 1997, 14, 939-944.	1.3	3
149	CRAFT: Compact genome Representation toward large-scale Alignment-Free daTabase. Bioinformatics, 2021, 37, 155-161.	4.1	3
150	Abstract 24: Multi-feature ensemble learning on cell-free dna for accurately detecting and locating cancer. Cancer Research, 2021, 81, 24-24.	0.9	3
151	A model-based approach to assess reproducibility for large-scale high-throughput MRI-based studies. NeuroImage, 2022, 255, 119166.	4.2	3
152	Phageâ€“bacterial contig association prediction with a convolutional neural network. Bioinformatics, 2022, 38, i45-i52.	4.1	3
153	Computational methods for the analysis of tag sequences in metagenomics studies. Frontiers in Bioscience - Scholar, 2012, S4, 1333-1343.	2.1	2
154	Marine Bacterial, Archaeal, and Protistan Association Networks. , 2013, , 1-10.		2
155	Alignment-Free Sequence Comparison Based on Next Generation Sequencing Reads: Extended Abstract. Lecture Notes in Computer Science, 2012, , 272-285.	1.3	2
156	Chromatin Regulation and Gene Centrality Are Essential for Controlling Fitness Pleiotropy in Yeast. PLoS ONE, 2009, 4, e8086.	2.5	2
157	MLR-OOD: A Markov Chain Based Likelihood Ratio Method for Out-Of-Distribution Detection of Genomic Sequences. Journal of Molecular Biology, 2022, 434, 167586.	4.2	2
158	HiFine: integrating Hi-c-based and shotgun-based methods to reFine binning of metagenomic contigs. Bioinformatics, 2022, , .	4.1	2
159	Whole Genome Amplification and Branching Processes. Advances in Applied Probability, 1997, 29, 629-668.	0.7	1
160	Sequential sib-pair and association studies to detect genes in quantitative traits. Genetic Epidemiology, 1997, 14, 885-890.	1.3	1
161	Usefulness and limitations of dK random graph models to predict interactions and functional homogeneity in biological networks under a pseudo-likelihood parameter estimation approach. BMC Bioinformatics, 2009, 10, 277.	2.6	1
162	Meeting report on RECOMB 2013 (the 17th Annual International Conference on Research in) Tj ETQq0 0 0 rgBT /Overlock 10 Tf 50 62 T	0.5	1

#	Article	IF	Citations
163	Network tuned multiple rank aggregation and applications to gene ranking. BMC Bioinformatics, 2015, 16, S6.	2.6	1
164	Towards enhanced and interpretable clustering/classification in integrative genomics. Nucleic Acids Research, 2017, 45, e169-e169.	14.5	1
165	Confidence intervals for Markov chain transition probabilities based on next generation sequencing reads data. Quantitative Biology, 2020, 8, 143-154.	0.5	1
166	KIMI: Knockoff Inference for Motif Identification from molecular sequences with controlled false discovery rate. Bioinformatics, 2021, 37, 759-766.	4.1	1
167	Reduced fecal shortâ€chain fatty acids in hispanic children with ulcerative colitis. Physiological Reports, 2021, 9, e14918.	1.7	1
168	Sparse logistic regression revealed the associations between HBV PreS quasispecies and hepatocellular carcinoma. Virology Journal, 2022, 19, .	3.4	1
169	Pooling Strategies for Establishing Physical Genome Maps Using FISH. Journal of Computational Biology, 1997, 4, 467-486.	1.6	0
170	Using case-control designs for genome-wide screening for associations between genetic markers and disease susceptibility loci. Genetic Epidemiology, 1999, 17, S779-S784.	1.3	0
171	New Estimator of the Genotype Risk Ratio for Use in Case-Parental Control Studies. American Journal of Epidemiology, 2001, 154, 259-263.	3.4	0
172	Data integration for the study of protein interactions. Chapman & Hall/CRC Mathematical and Computational Biology Series, 2009, , 259-274.	0.1	0
173	Preface: Special Issue. Journal of Computational Biology, 2012, 19, 575-576.	1.6	0
174	Preface: Research in Computational Molecular Biology (RECOMB 2013). Journal of Computational Biology, 2013, 20, 713-713.	1.6	0
175	International Workshop on Applications of Probability and Statistics to Biology, July 11â€â€13, 2019 â€“â€“In Honor of Professor Minping Qianâ€™s 80th Birthday. Quantitative Biology, 2020, 8, 177-186.	0.5	0
176	Dynamic Programming Algorithms for Haplotype Block Partitioning and Tag SNP Selection Using Haplotype Data or GenotypeÂData. Lecture Notes in Computer Science, 2004, , 96-112.	1.3	0
177	Computational Methods for Predicting Domainâ€“Domain Interactions. Computational Biology, 2010, , 157-173.	0.2	0
178	Predicting kinase functional sites using hierarchical stochastic language modelling. Statistics and Its Interface, 2010, 3, 523-531.	0.3	0
179	Inferring Signaling and Gene Regulatory Network from Genetic and Genomic Information. , 2011, , 479-500.		0
180	Accurate Genome Relative Abundance Estimation Based on Shotgun Metagenomic Reads. , 2014, , 1-5.		0

#	Article	IF	Citations
181	Extended Local Similarity Analysis (eLSA) of Biological Data. , 2014, , 1-5.		0
182	Statistical Methods for Pairwise Comparison of Metagenomic Samples. Frontiers in Probability and the Statistical Sciences, 2021, , 81-99.	0.1	0
183	An Efficient Algorithm for Deciphering Regulatory Motifs. , 2007, , 249-269.		0
184	A New Context Tree Inference Algorithm for Variable Length Markov Chain Model with Applications to Biological Sequence Analyses. Journal of Computational Biology, 2022, , .	1.6	0
185	Michael Waterman's Contributions to Computational Biology and Bioinformatics. Journal of Computational Biology, 0, , .	1.6	0